CCCGGGCCCGGCCGCCGTCGCGGGACCCCGTGCCCGGCCGCCGTCGCCACCGCCGCCCCGGCCGACCGAG
GGACCCGCCCGCCCGCGGCTGCTCCGGACCTAGAGGATCAAGTCATAATGGGAGCATTTTTAGACAAGCC
AAAGATGGAGAAGCATAATGCCCAGGGGCAGGGGAATGGGTTACGATACGGCCTAAGCAGCATGCAAGGT
TGGCGAGTTGAAATGGAGGACGCACACACGGCTGTGATCGGTTTGCCAAGTGGACTTGAGACATGGTCAT
TCTTTGCTGTATATGATGGGCATGCTGGTTCTCAGGTTGCCAAATACTGCTGTGAGCACTTGTTAGATCA
CATCACCAATAACCAGGATTTCAGAGGATCTGCAGGAGCACCTTCTGTGGAGAACGTAAAGAATGGAATC
AGAACAGGGTTTCTGGAGATTGATGAACACATGAGAGTTATGTCAGAGAAGAAACATGGTGCAGATAGAA
GCGGGTCAACAGCTGTGGGCGTCTTAATCTCTCCCCAACATACTTATTTCATTAACTGTGGAGACTCGAG
AGGTTTACTTTGTAGGAATAGAAAAGTTCACTTCTTCACACAAGACCATAAACCAAGTAACCCGCTGGAA
AAAGAACGAATTCAGAATGCAGGGGGCTCGGTGATGATTCAGCGTGTCAATGGCTCTCTGGCTGTATCGA
GGGCCCTTGGGGATTTCGATTACAAATGTGTCCATGGAAAAGGTCCCACAGAGCAGCTTGTCTCCCCAGA
GCCCGAAGTCCATGATATTGAAAGGTCTGAAGAAGATGACCAGTTCATCATCCTTGCATGCGATGGCATC
TGGGACGTCATGGGGAACGAAGAGCTCTGTGACTTTGTGAGATCCAGACTTGAAGTCACTGATGACCTTG
AGAAAGTTTGCAATGAAGTAGTCGACACCTGCTTGTATAAGGGAAGTCGAGACAACATGAGTGTGATTTT
GATCTGTTTTCCAAGTGCACCCAAAGTCTCGGCAGAGGCGGTGAAGAAGGAGGCGGAGCTGGACAAGTAC
CTGGAGAGCAGAGTAGAAGAAATCATAAAGAAGCAGGTGGAAGGCGTCCCTGACTTAGTCCACGTGATGC
GCACGTTAGCCAGTGAGAACATCCCCAGCCTCCCACCAGGGGGTGAATTGGCAAGCAAGCGGAATGTAAT
TGAAGCCGTTTACAATAGACTGAACCCTTACAAAAATGACGACACTGATTCTGCGTCAACCGATGATATG
TGGTAAAGCCGCTCACCCAGCCGTGGACTCACCTTCGCCTGCAAAGGGGAAGCCAGCTCATCCTTGCCGA
GCCTTTACCATCCATCACCGACTTCACAGGAGGGTCTGACACGGGTGAGGACTGCAG (SEQ NO:1)

MGAFLDKPKMEKHNAQGQGNGLRYGLSSMQGWRVEMEDAHTAVIGLPSGLE
TWSFFAVYDGHAGSQVAKYCCEHLLDHITNNQDFRGSAGAPSVENVKNGIRTG
FLEIDEHMRVMSEKKHGADRSGSTAVGVLISPQHTYFINCGDSRGLLCRNRKVH
FFTQDHKPSNPLEKERIQNAGGSVMIQRVNGSLAVSRALGDFDYKCVHGKGPT
EQLVSPEPEVHDIERSEEDDQFIILACDGIWDVMGNEELCDFVRSRLEVTDDLEK
VCNEVVDTCLYKGSRDNMSVILICFPSAPKVSAEAVKKEAELDKYLESRVEEIIK
KQVEGVPDLVHVMRTLASENIPSLPPGGELASKRNVIEAVYNRLNPYKNDDTDS
ASTDDMW (SEQ ID NO:2)

FIGURE 1

underlined = deleted in targeting construct

[   ] = sequence flanking Neo insert in targeting construct

CCCGGGCCCGGCCGCCGTCGCGGGACCCCGTGCCCGGCCGCCGTCGCCACCGCCGCCCCG
GCCGACCGAGGGACCCGCCCGCCCGCGGCTGCTCCGG[ACCTAGAGGATCAAGTCATAATG
GGAGCATTTTTAGACAAGCCAAAGATGGAGAAGCATAATGCCC]**AGGGGCAGGGGAATGGG**
TTACGATACGGCCTAAGCAGCATGCAAGGTTGGCGAGTTGAAATGGAGGACGCACACACG
GCTGTGATCGGTTTGCCAAGTGGACTTGAGACATGGTCATTCTTTGCTGTATATGATGGG
CATGCTGGTTCTCAGGTTGCCAAATA[CTGCTGTGAGCACTTGTTAGATCACATCACCAAT
AACCAGGATTTCAGAGGATCTGCAGGAGCACCTTCTGTGGAGAACGTAAAGAATGGAATC
AGAACAGGGTTTCTGGAGATTGATGAACACATGAGAGTTATGTCAGAGAAGAAACATGGT
GCAGATAGAAGCGGGTCAACAGCTGTGGGCGTCTTAATCTCTCCCCAACATACTTATTTC
ATTAACTGTGGAGACTCGAGAGGTTTACTTTGTAGGAATAGAAAAGTTCACTTCTTCACA
CAAGACCATAAACCAAGTAACCCGCTGGAAAAAGAACGAATTCAGAATGCAGGGGGCTCG
GTGATGATTCAGCGTGTCAATGGCTCTCTGGCTGTATCGAGGGCCCTTGGGGATTTCGAT
TACAAATGTGTCCATGG]AAAAGGTCCCACAGAGCAGCTTGTCTCCCCAGAGCCCGAAGTC
CATGATATTGAAAGGTCTGAAGAAGATGACCAGTTCATCATCCTTGCATGCGATGGCATC
TGGGACGTCATGGGGAACGAAGAGCTCTGTGACTTTGTGAGATCCAGACTTGAAGTCACT
GATGACCTTGAGAAAGTTTGCAATGAAGTAGTCGACACCTGCTTGTATAAGGGAAGTCGA
GACAACATGAGTGTGATTTTGATCTGTTTTCCAAGTGCACCCAAAGTCTCGGCAGAGGCG
GTGAAGAAGGAGGCGGAGCTGGACAAGTACCTGGAGAGCAGAGTAGAAGAAATCATAAAG
AAGCAGGTGGAAGGCGTCCCTGACTTAGTCCACGTGATGCGCACGTTAGCCAGTGAGAAC
ATCCCCAGCCTCCCACCAGGGGGTGAATTGGCAAGCAAGCGGAATGTAATTGAAGCCGTT
TACAATAGACTGAACCCTTACAAAAATGACGACACTGATTCTGCGTCAACCGATGATATG
TGGTAAAGCCGCTCACCCAGCCGTGGACTCACCTTCGCCTGCAAAGGGGAAGCCAGCTCA
TCCTTGCCGAGCCTTTACCATCCATCACCGACTTCACAGGAGGGTCTGACACGGGTGAGG
ACTGCAG

FIGURE 2A

Gene Sequence Structure *

164 bp Sequence Deleted 326 bp
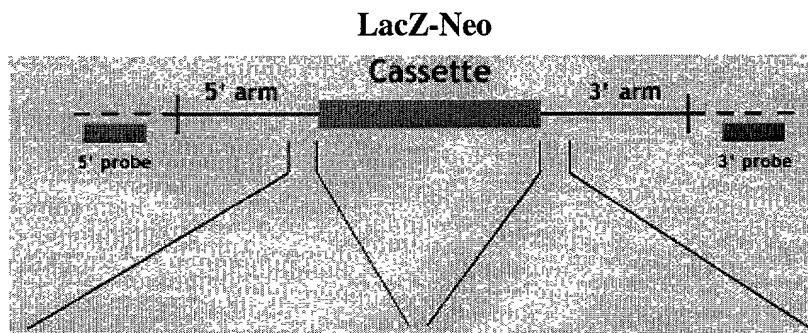
Size of full-length cDNA: 1387 bp

Targeting Vector* (genomic sequence)

Construct Number: 2639

Arm Length:
5': 2.5 kb
3': 4 kb

LacZ-Neo
Cassette

5' arm   3' arm

5' probe   3' probe

————— Targeting Vector
– – – – – Endogenous Locus

* Not drawn to scale

5'>GCAACACAATGCTTGTAGGTA
TAGCCTGTGAGTTTTTCCAGCTTC
CTTGTATCTTATAGATTCTGGGTA
AAGAGTGTTGGACATGTTTTGTTT
CAAAGGCAATCACTTATTTTCTTA
TTTCTCTTCCTTTACAGACCTAGA
GGATCAAGTCATAATGGGAGCATT
TTTAGACAAGCCAAAGATGGAGAA
GCATAATGCCC<3' (SEQ ID
NO:3)

5'>CTGCTGTGAGCACTTGTTAGA
TCACATCACCAATAACCAGGATTT
CAGAGGATCTGCAGGAGCACCTTC
TGTGGAGAACGTAAAGAATGGAAT
CAGAACAGGGTTTCTGGAGATTGA
TGAACACATGAGAGTTATGTCAGA
GAAGAAACATGGTGCAGATAGAAG
CGGGTCAACAGCTGTGGGCGTCTT
AATCTCTCCCC<3'(SEQ ID
NO:4)

FIGURE 2B